MA 3046 - Matrix Analysis
Laboratory Number 5
Least Squares Solutions and Numerical Accuracy

As discussed in class, when a given system of equations

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}$$

involves more equations than unknowns, i.e. when $m > n$, or when the system is *overdetermined*, or when, more colloquially, the matrix $\mathbf{A}$ is "tall and skinny", the system is generally *not* solvable in the "classical" (Gaussian elimination or equality) sense. (The exception occurs in the unusual instance where $\mathbf{b}$ just happens to lie in Col($\mathbf{A}$).)

In this situation, the best that can generally can be expected is to find an $\mathbf{x}$ that will "closely" solve the system, in the sense that the values of $\mathbf{A}\,\mathbf{x}$ computed on the left using that $\mathbf{x}$ will closely approximate those of $\mathbf{b}$ on the right. The only drawback with this approach is that the particular $\mathbf{x}$ that one finds generally depends on how they wish to measure "close," and there is more than one reasonable way to do this. The arguably most common measure, however, is the Euclidean norm of the residual,

$$\|\,\mathbf{r}\,\|_2 \;=\; \|\,\mathbf{b} - \mathbf{A}\,\mathbf{x}\,\|_2 \;\;,$$

i.e. precisely the distance from $\mathbf{A}\,\mathbf{x}$ to $\mathbf{b}$. The minimum value of this quantity occurs when $\mathbf{r}$ is orthogonal to Col($\mathbf{A}$), or, equivalently, when $\mathbf{r}$ is orthogonal either to each column of $\mathbf{A}$ (or to each row of $\mathbf{A}^H$). This last condition, however, expressed in matrix form, becomes

$$\mathbf{A}^H\,(\mathbf{b} - \mathbf{A}\,\mathbf{x}) \;=\; \mathbf{0} \qquad \Longrightarrow \qquad \mathbf{A}^H\,\mathbf{A}\,\mathbf{x} \;=\; \mathbf{A}^H\,\mathbf{b} \tag{1}$$

This form is commonly called the *normal equations.*

We can easily show that Null($\mathbf{A}$) is equal to Null($\mathbf{A}^H\,\mathbf{A}$), and therefore the normal equations will be uniquely solvable if and only if Null($\mathbf{A}$)$= \{\mathbf{0}\}$. The entire situation, however, becomes a bit muddier when we must factor in the effects of floating-point arithmetic. Specifically, we can show that, in general, the singular values of $\mathbf{A}^H\,\mathbf{A}$ are precisely the squares of the singular value of $\mathbf{A}$, and therefore, in the Euclidean norm:

$$\text{cond}\left(\mathbf{A}^H\,\mathbf{A}\right) \;=\; \left(\text{cond}\,(\mathbf{A})\right)^2$$

Therefore matrices that are mildly ill-conditioned, i.e. for which cond($\mathbf{A}$) $\doteq eps^{-1/2}$, may not be accurately solvable when converted to the normal equations form. Fortunately, the MATLAB backslash function automatically takes care of all these issues, at least as well any numerical method can. Specifically, it uses basic Gaussian elimination, implemented via $\mathbf{L}\,\mathbf{U}$ decomposition, when $\mathbf{A}$ is square. However, when $m > n$, backslash switches to least squares, although implemented by a method, apparently related to the $\mathbf{QR}$ factorization, which is equivalent to the normal equations, but numerically preferable.

However, when $\mathbf{A}$ has a fairly large condition number, even MATLAB can have difficulties. Before implementing least squares in such situations, it is important to recall that least squares in effect projects $\mathbf{b}$ onto $\mathrm{Col}(\mathbf{A})$, and that ill-conditioning generally reflects the fact that one or more columns of $\mathbf{A}$ are linearly dependent. Therefore, if we could simply remove those columns, we would be left with ones that were "really" linearly independent, but spanned essentially the same column space, and which should hence produce essentially the same projection. However, if properly chosen, this reduced set of columns should yield a reasonably well-conditioned matrix. The only practical problem becomes deciding which columns could be removed without significantly affecting the column space. But given either the SVD or the $\mathbf{QR}$ factorization, such a choice should not be too difficult.

Name: ____________________

MA 3046 - Matrix Analysis
Laboratory Number 5
Least Squares Solutions and Numerical Accuracy

1. Switch to the your lab directory and start MATLAB. Download the file

least_sq_data.mat

to your disk.

2. Give the command

load least_sq_data

to input the random 20×4 matrix $\mathbf{a}$, and determine its condition number. (Save this matrix and also print it out.)

$$\text{cond}(\mathbf{a}) = \underline{\hspace{3cm}}$$

Then create the four-element column vector **xexact**

$$\mathbf{xexact} = \begin{bmatrix} 2 \\ -3 \\ 1 \\ 0 \end{bmatrix}$$

Finally, create a twenty-element column vector $\mathbf{b}$ by giving the MATLAB command

$$\mathbf{b} = \mathbf{a} * \mathbf{xexact}$$

In what subspace of $\mathbb{R}^{20}$ should $\mathbf{b}$ lie?

3. Briefly explain what should happen if you were to apply Gaussian elimination to the system:

$$\mathbf{a}\,\mathbf{x} = \mathbf{b}$$

for this **a** and **b**.

4. Create and store as **qhat** and **rhat**, respectively, the matrices $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$ associated with the *reduced* **QR** factorization of the matrix **a** created in part 2. Record the values of

$$\mathbf{rhat} = \left[\qquad\qquad\qquad \right]$$

Briefly describe why the values in **rhat** either do or do not seem reasonable in light of the value of condition number of **a** as determined in that same part.

5. Compute the singular values of the matrix **a** created in part 2.

Briefly describe why these values either do or do not seem reasonable in light of the value of condition number of **a** as determined in that same part.

6. Compute

$$\mathbf{xbs} = \mathbf{a} \backslash \mathbf{b} \quad , \quad \mathbf{xne} = (\mathbf{a}' * \mathbf{a}) \backslash (\mathbf{a}' * \mathbf{b}) \quad \text{and} \quad \mathbf{xqr} = \mathbf{rhat} \backslash \left(\mathbf{qhat}' * \mathbf{b}\right)$$

and record their values:

$$\mathbf{xbs} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \end{bmatrix} \quad , \quad \mathbf{xne} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \end{bmatrix} \quad \text{and} \quad \mathbf{xqr} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \end{bmatrix}$$

What should these be computing?

xbs –

xne –

xqr –

Briefly explain why these values appear either reasonable or unreasonable, considering the value of **xexact** as created in part 2.

7. Find the norms:

$$\| \mathbf{xbs} - \mathbf{xne} \|_2 \;\; -$$

$$\| \mathbf{xbs} - \mathbf{xqr} \|_2 \;\; -$$

What do these values say about solving least squares and about MATLAB's backslash algorithm when $m > n$ and the matrix is well-conditioned?

8. Next form the residual

$$\mathbf{r} = \mathbf{b} - \mathbf{a}\text{*}\mathbf{xbs}$$

How does this result agree or not agree with the way you formed **b** in part 2 above.

Similarly analyze

$$\mathbf{r} = \mathbf{b} - \mathbf{a}*\mathbf{xne} \quad \text{and} \quad \mathbf{r} = \mathbf{b} - \mathbf{a}*\mathbf{xqr}$$

9. Now, create a column vector **bp** by giving the MATLAB command

$$\mathbf{bp} = \mathbf{b} + \mathbf{0.05}* \ \mathbf{randn(20,1)}$$

Briefly describe what this should do?

10. Now compute both

$$\mathbf{xpbs} = \mathbf{a} \backslash \mathbf{bp} \ \ , \ \ \mathbf{xpne} = (\mathbf{a}' * \mathbf{a}) \backslash (\mathbf{a}' * \mathbf{bp}) \ \text{ and } \ \mathbf{xpqr} = \mathbf{rhat} \backslash (\mathbf{qhat}' * \mathbf{bp})$$

and record their values:

$$\mathbf{xpbs} = \begin{bmatrix} \quad & \quad \\ & \\ & \end{bmatrix} \ , \ \ \mathbf{xpne} = \begin{bmatrix} \quad & \quad \\ & \\ & \end{bmatrix} \ \text{ and } \ \mathbf{xpqr} = \begin{bmatrix} \quad & \quad \\ & \\ & \end{bmatrix}$$

Briefly explain why these values either are or are not reasonable considering the values you computed in part 7 and your answer to part 9 above.

11. Find the norms:

$$\| \mathbf{xpbs} - \mathbf{xpne} \|_2 \;\; -$$

$$\| \mathbf{xpbs} - \mathbf{xpqr} \|_2 \;\; -$$

Briefly describe whether these values still support your conclusion in part 7 about MATLAB's backslash algorithm when $m > n$ and the matrix is well-conditioned?

12. Next form the residual

$$\mathbf{rp} = \mathbf{bp} - \mathbf{a*xpbs}$$

and contrast the value here with the value of the residual computed in part 8. Briefly explain why or why not any differences appear reasonable.

Now compute $\quad \mathbf{a}' {*} \mathbf{rp} = \begin{bmatrix} \\ \\ \\ \qquad\qquad \\ \\ \\ \end{bmatrix} =$

Briefly describe how well this result does or does not agree with theory?

13. Also create, by hand, the upper triangular matrix:

$$\mathbf{u} = \begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 10^{-8} \end{bmatrix}$$

and the matrix

$$\mathbf{aa} = \mathbf{a} * \mathbf{u}$$

What are the MATLAB rank and condition number of **aa**?

$$\text{rank}(\mathbf{aa}) = \underline{\hspace{4cm}}$$

$$\text{cond}(\mathbf{aa}) = \underline{\hspace{4cm}}$$

Which columns of **a** should approximately span Col(**aa**)?

14. Compute the singular values of the matrix **aa** created in part 13.

Briefly describe why these values either do or do not seem reasonable in light of the way in which **aa** was constructed, and the value of condition number of **aa** as determined in that same part.

15. Create and store as **qqhat** and **rrhat**, respectively, the matrices associated with the *reduced* **QR** factorization of the matrix **aa** created in part 13. Record the values of

$$\mathbf{rrhat} = \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

Briefly describe why the values in **rrhat** either do or do not seem reasonable in light of your answers to the parts 13 and 14.

16. Now compute both

$$\mathbf{xxbs} = \mathbf{aa} \backslash \mathbf{bp} \quad , \quad \mathbf{xxne} = (\mathbf{aa}' * \mathbf{aa}) \backslash (\mathbf{aa}' * \mathbf{bp})$$

and

$$\mathbf{xxqr} = \mathbf{rrhat} \backslash (\mathbf{qqhat}' * \mathbf{bp})$$

and record their values:

$$\mathbf{xxbs} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \end{bmatrix} , \quad \mathbf{xxne} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \end{bmatrix}$$

and

$$\mathbf{xxqr} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \end{bmatrix}$$

What should these vectors represent and what do your actual values say about MATLAB's backslash algorithm and the normal equations when $m > n$ and the matrix is not well-conditioned?

17. Create the matrix

$$\mathbf{aa3} = \mathbf{aa(\ : \ , \ 1{:}3 \ )}$$

What are the MATLAB rank and condition number of **aa3**?

$$\text{rank}(\mathbf{aa3}) = \underline{\hspace{4cm}}$$

$$\text{cond}(\mathbf{aa3}) = \underline{\hspace{4cm}}$$

18. Now compute both

$$\mathbf{xx3bp} = \mathbf{aa3} \backslash \mathbf{bp} \quad \text{and} \quad \mathbf{xx3ne} = (\mathbf{aa3'} * \mathbf{aa3}) \backslash (\mathbf{aa3'} * \mathbf{bp})$$

and record their values:

$$\mathbf{xx3bp} = \begin{bmatrix} \hspace{3cm} \\ \\ \\ \\ \end{bmatrix} \qquad \mathbf{xx3ne} = \begin{bmatrix} \hspace{3cm} \\ \\ \\ \\ \end{bmatrix}$$

What should these now represent? Compare these values both to each other and to the values you computed in part 16. What conclusion can you draw from that comparison?

19. Now form the residuals

$$\mathbf{r1} = \mathbf{bp} - \mathbf{aa}{*}\mathbf{xxbs}\ , \quad \mathbf{r2} = \mathbf{bp} - \mathbf{aa}{*}\mathbf{xxne}\ , \quad \mathbf{r3} = \mathbf{bp} - \mathbf{aa3}{*}\mathbf{xx3bp}$$

and

$$\mathbf{r4} = \mathbf{bp} - \mathbf{aa3}{*}\mathbf{xx3ne}$$

and compute their Euclidean norms:

$$\| \mathbf{r1} \|_2 =$$

$$\| \mathbf{r2} \|_2 =$$

$$\| \mathbf{r3} \|_2 =$$

$$\| \mathbf{r4} \|_2 =$$

What do these values and the values of the various $\mathbf{x}$'s computed in parts 16-18 say about the advisability and costs, if any, of removing nearly-linearly dependent columns from ill-conditioned systems?